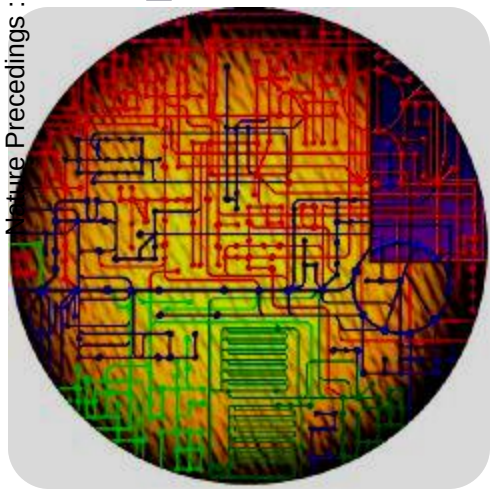
Computational systems biology

Impact on drug discovery



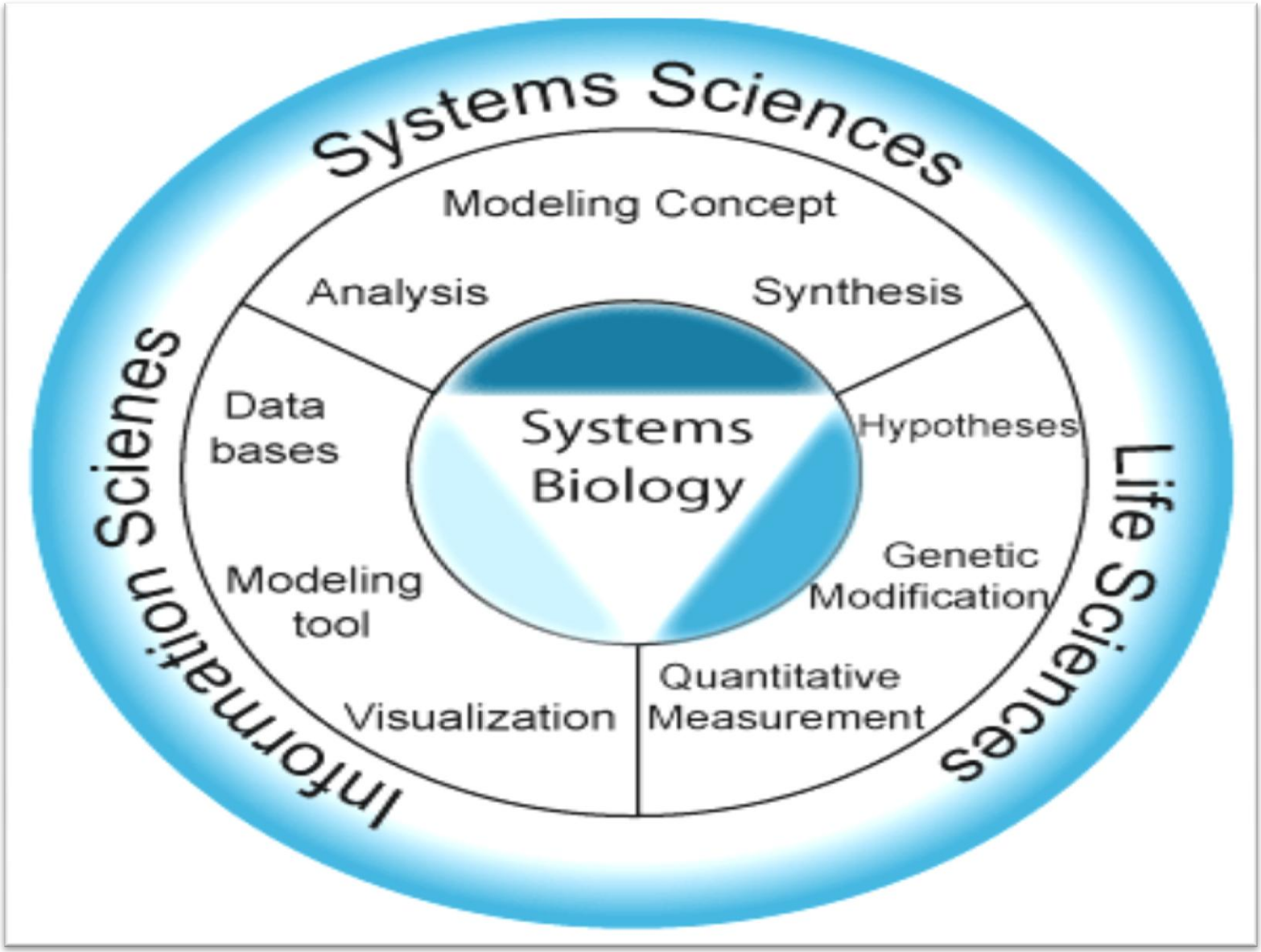
Guide by
M. Elizabeth Sobhia
Assistant Professor
mesophia@niper.ac.in

Presented by
Ankit Geete
2nd semester
Dept. of Pharmacoinformatics
NIPER, Mohali

Definition

- Systems biology is a multi-disciplinary field that deals with mechanisms involved in complex biological processes by considering them as integrated systems of multiple interacting components.

- The huge amount of data involved in this study necessitates the use of computational tools



General process in systems biology

- Collecting large sets of experimental data
- Genomic, Proteomic or Metabolomic data
- Generating predictive models of those data using computers
- Assessing or correcting those computer models
- Comparing the predicted data with newly derived experimental data

Descriptive science vs. Predictive science

Then why we chose it……..

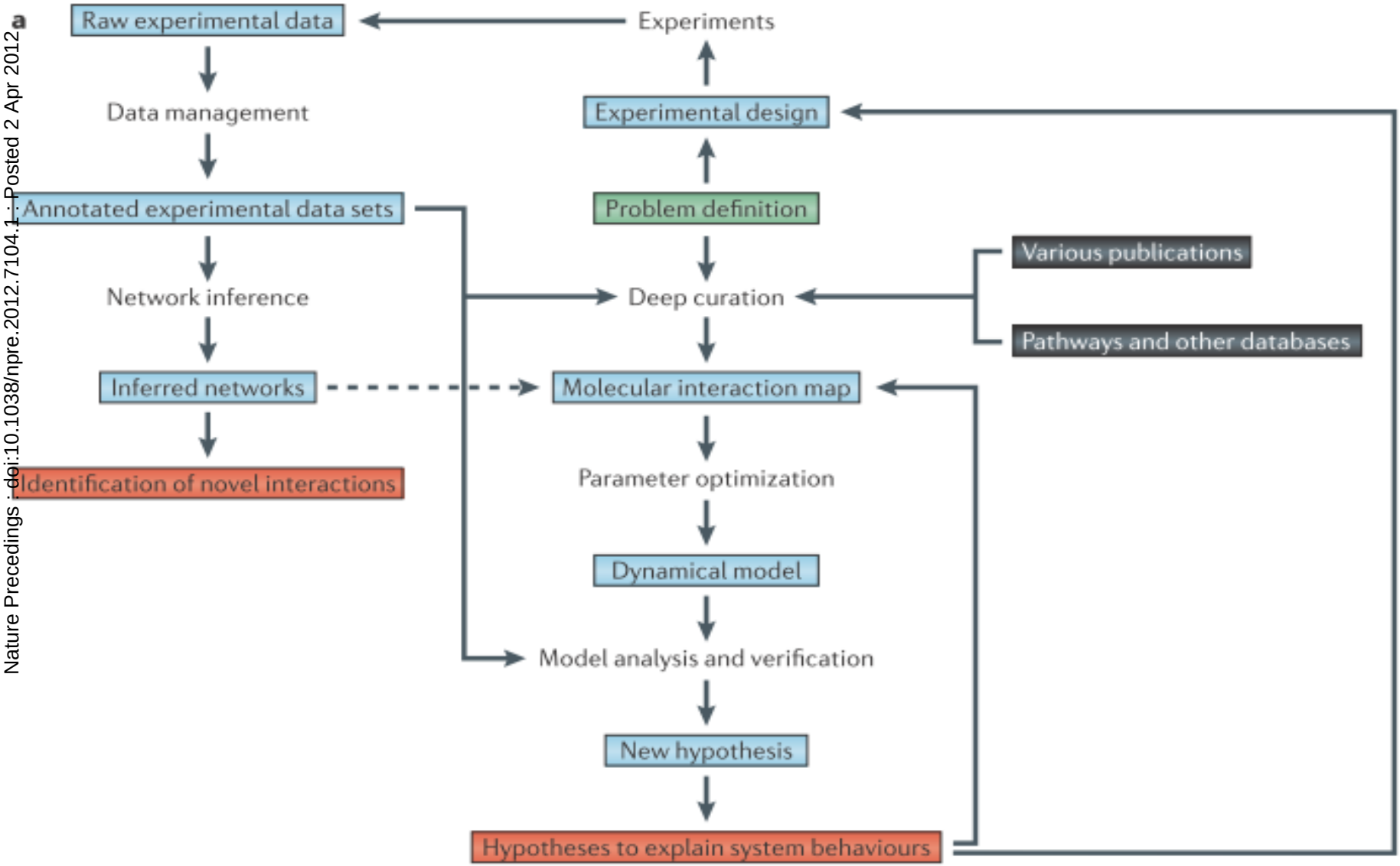
- Produce useful predictions or extrapolations that match experimental results
- Beyond present-day experimental capabilities
- Save time, cost or effort
- Identify missing components, processes or functions in a system
- Enable complex processes to be better understood or visualized

Software for systems biology

- Computational tools are needed at each step of computational systems biology workflow. These are –

  - **Data handling**
  - **Network inference**
  - **Deep curation**
  - **Dynamical simulation**
  - **Model analysis**

Workflow of computational tasks in systems biology



Nature Reviews Genetics, 12 (December 2011)

Software for systems biology

- **Data handling**
- Network inference
- Deep curation
- Dynamical simulation
- Model analysis

Data management

- Proper acquisition and handling of data is crucially important for both the generation and verification of hypotheses.

Data-management standards

- Standards for data management have focused on three core aspects:
  - **Minimum information**
  - **File formats**
  - **Ontologies**

Minimum information

Minimum information is a checklist of required supporting information for data sets from different experiments

HUPO
Human Proteome Organisation

FGED SOCIETY
Functional Genomics Data Society

Organization define Minimum information

Human Proteome Organisation

Minimum Information About a Proteomic Experiment (MIAPE)

- Human Proteome Project (HPP)
- Human Plasma Proteome Project (HPPP)
- Human Liver Proteome Project (HLPP)
- Human Brain Proteome Project (HBPP)
- Human Antibody Initiative (HAI)
- Proteomics Standards Initiative (PSI)
  - People and Funding
  - Visit web site
- Human Disease Glycomics / Proteome Initiative (HGPI)
- Human Kidney and Urine Proteome Project (HKUPP)
- Mouse Models of Human Disease
- HUPO CardioVascular Initiative (HCVI)
- Proteome Biology of Stem Cells Initiative
- Disease Biomarkers Initiatives

Proteomics Standards Initiative

Overview of Project

The HUPO Proteomics Standards Initiative (PSI) defines community standards for data representation in proteomics to facilitate data comparison, exchange and verification.

The main organizational unit of the Proteomics Standards Initiative is the work group. Currently, there are the following work groups:.

- Molecular Interactions (MI)
- Mass Spectrometry (MS)
- Proteomics Informatics (PI)
- Protein Modifications (MOD)
- Protein Separation (PS)

The standard deliverables of each work group are

- Minimum Information Specification: For the given domain, this specifies the minimum information required for the useful reporting of experimental results in this domain.
- Formal exchange format for experimental results in the domain. This will usually be an XML format, capable of representing at least the Minimum Information, and normally significant additional detail.
- Controlled vocabularies.
- Support for implementation of the standard in publicly available tools.

Collaborations with other Initiatives

The HUPO PSI has a subsidiary role supporting other HUPO projects and workgroups, aiming to ensure collaboration and consistency in the data management approaches taken by the HUPO proteomics initiatives. We continued to data management and long-term data repository support for the HUPO Plasma Proteome (I), Brain and Liver Proteome Projects. Currently, major proteomics repositories represented in the PSI mass spectrometry work group (PRIDE,

Organization define Minimum information

msi-Metabolomics Standards Initiative

MSI WGs
- Biological context metadata WG
- Chemical analysis WG
- Data processing WG
- Ontology WG
- Exchange format WG

DOCUMENTATION
- Minutes
- Presentations

MEETINGS
- MSI@PSI

LINKS
- Metabolomics Society
- Conferences

The Metabolomics Standards Initiative (MSI)

The Metabolomics Society has appointed an Oversight Committee to monitor, coordinate and review the efforts of working groups (WG) in specialist areas that will examine standardization and make recommendations. The overall chair of this committee is Oliver Fiehn. The five MSI WGs, some of which are divided into further subgroups, are listed here:

- Biological context metadata WG
- Chemical analysis WG
- Data processing WG
- Ontology WG
- Exchange format WG

The structure of the WGs thus follows the general "workflow" model in metabolomics: from a description of the study design to sample workup, data acquisition, processing and export, bound together by controlled vocabularies and relationships between the terms used.

The MSI WGs are working towards developing the following standards:

1. **Core Information for Metabolomics Reporting (CIMR):** This document will specify the minimal guidelines reporting metabolomics work. It will do so in a textual form and will seek in the long term to cover all application areas and analysis technologies. This document will be developed by the biological context metadata WG, the chemical analysis WG, the data processing WG, the exchange format WG and the ontology WG.

2. **Ontology:** The CV and ontology will be developed iteratively by the ontology WG on the basis of CIMR. This will be done in collaboration with: HUPO-PSI, MGED and FuGO.

3. **Exchange format:** A data model and exchange format will be developed by the exchange format WG on the

Organization define Minimum information

FGED SOCIETY

Minimum Information About a Microarray Experiment (MIAME)

HOME | MEETINGS | WORKGROUP

FGED

The Functional Genomics Data Society - FGED Society, founded in 1999 as the MGED Society, advocates for open access to genomic data sets and works towards providing concrete solutions to achieve this. Our goal is to assure that investment in functional genomics data generates the maximum public benefit. Our work on defining minimum information specifications for reporting data in functional genomics papers have already enabled large data sets to be used and reused to their greater potential in biological and medical research.

We work with other organisations to develop standards for biological research data quality, annotation and exchange. We facilitate the creation and use of software tools that build on these standards and allow researchers to annotate and share their data easily. We promote scientific discovery that is driven by genome wide and other biological research data integration and meta-analysis.

As part of this effort, we are providing a service for facilitating data deposition.

Defined FGED Standards:

MIAME MINSEQE MAGE-TAB MAGE MGED Ontology

Collaborative Standards Working Groups:

MIBBI ISA-TAB FuGE OBI under OBO Foundry

Read more about: FGED Sponsors ▪ Historical highlights ▪ FGED Meetings ▪ FGED Supported Meetings ▪ FGED Workgroups ▪ Relevant Publications ▪ FGED Board

FGED Sponsors

illumina

Metadata bases

- Metadata (that is, data about data), which has led to the definition of standards such as

  - **International Organization for Standardization metadata registry (ISO–MDR) standard**

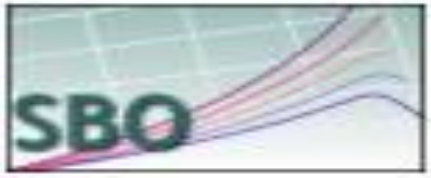
  - **Dublin Core Metadata Initiative (DCMI) standard**

Ontologies

- Ontologies define the relationships and hierarchy between different terms and allow the unique, semantic annotation of data

  - **Gene Ontology (GO)**

  - **Systems Biology Ontology (SBO)**



Data-management and data-analysis tools

Current data management systems can be broadly classified as –

- Spreadsheet based
- Web based
- Workflow management systems (WMSs)

Spreadsheet based

- **Pros -**
  - Most popular mode of data storage and communication in the life science community
  - Ease of use and sharing
- **Cons -**
  - Standardized practice for filling the spreadsheet is required.
  - Not supported on all software platforms
- **Example -**
  - MAGE-TAB (a spreadsheet-based, MIAME-supportive format for microarray data)
  - Investigation–Study–Assay (ISA)-TAB formats

Web based

- Online wiki-based document and project management
- Provide security and privacy options for data protection
- Custom-built information systems → electronic lab notebooks (ELN)

Workflow management systems (WMSs)

- Power to integrate different tools and services in a computational pipeline

KNIME

KNIME /PRODUCTS /APPLICATIONS /PARTNERS /SERVICES /RESOURCES /COMPANY

Search this site

KNIME Team Space

KNIME Team Space allows you to share your workflows, data and metanodes with your team.

/ More information.

1 2 3 4 5 6

KNIME - Professional Open-Source Software

KNIME (Konstanz Information Miner) is a user-friendly and comprehensive open-source data integration, processing, analysis, and exploration platform. From day one, KNIME has been developed using rigorous software engineering practices and is used by professionals in both industry and academia in over 60 countries.

KNIME grows to accommodate your demand for data analytics. While KNIME Team Space suits small teams, the KNIME Server provides support for a full corporate setting including user authentication, remote execution, scheduling, SOA integration and a configurable web browser user interface. When combined with KNIME Cluster Execution the workflows can be sent to a compute cluster for parallel execution.

/ More information about KNIME.

Gartner | 2010 COOL VENDOR

KNIME has been selected by Gartner as Cool Vendor 2010 in the key technology areas Analytics, Business Intelligence, and Performance Management. Details on the Gartner Cool Vendor program can be found here.

Forum & Documentation

KNIME tech

News

KNIME 2012 User Group Meeting 2012/02/21

KNIME Version 2.5.2 released 2012/02/15

KNIME Version 2.5.1 released & upcoming User Group Meeting in Zurich 2011/12/21

more

Events

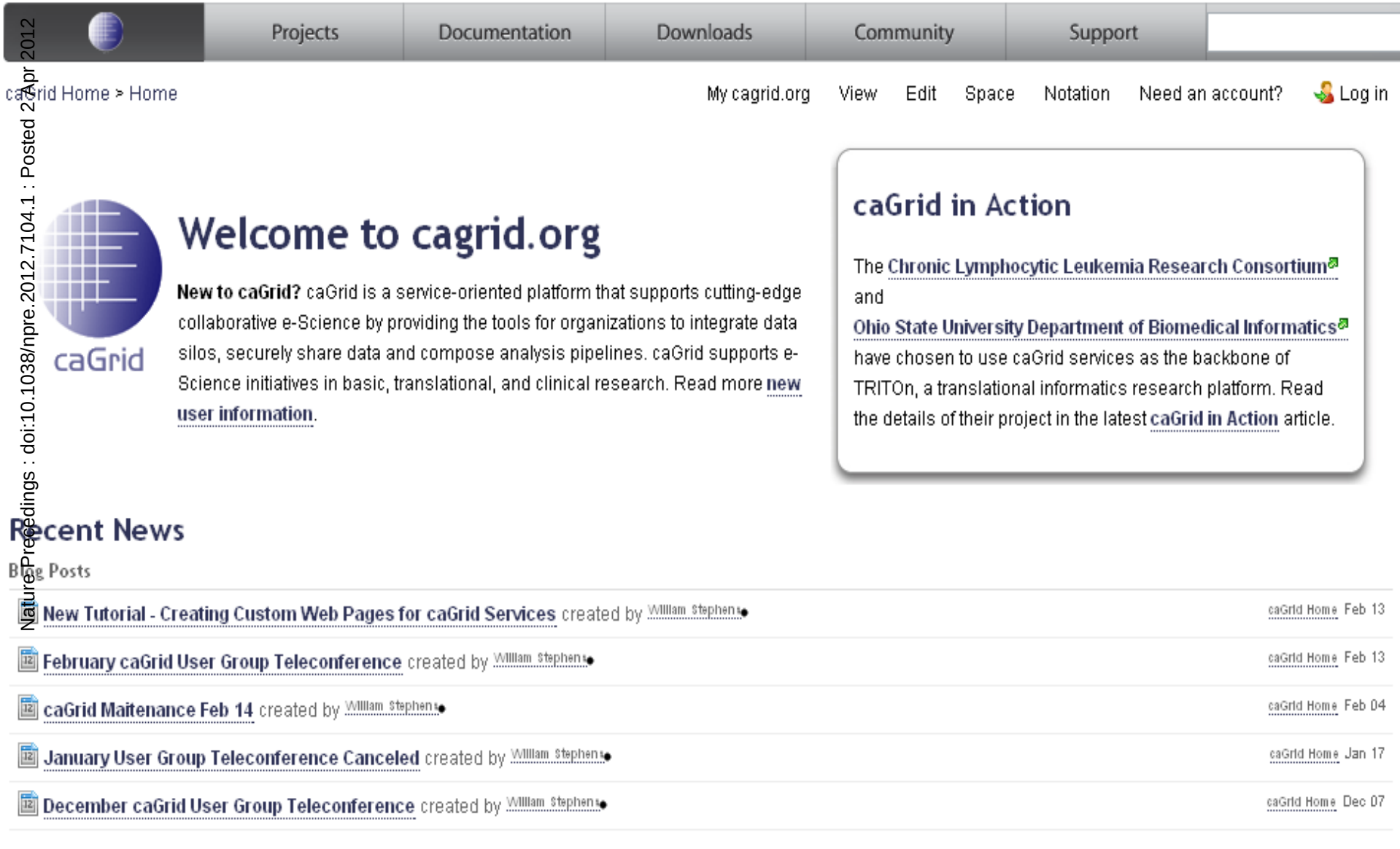
KNIME Basic and Advanced User Training - March 2012 2012/03/12 - 2012/03/14

KNIME Developer Training - April 2012 2012/04/16 - 2012/04/17

KNIME Basic and Advanced User Training - May 2012 2012/05/07 - 2012/05/09

more

caGrid



Taverna

Taverna

myGrid

Google™ Custom Search

Introduction Documentation Download Developers News Publications About

Taverna Workflow Management System

Powerful, scalable, open source & domain independent tools for designing and executing workflows. Access to 3500+ resources.

RECENT NEWS

- February 6, 2012 **SCUFL2: Taverna's new workflow format**
- January 23, 2012 **Software Sustainability Institute Collaborations Workshop**
- October 26, 2011 **Taverna/Galaxy next**

Get — Download for Windows, Mac OS X or Linux

Use — Learn about the features & functionality

Extend — Learn about the internals & how to develop plugins

COMMUNITY

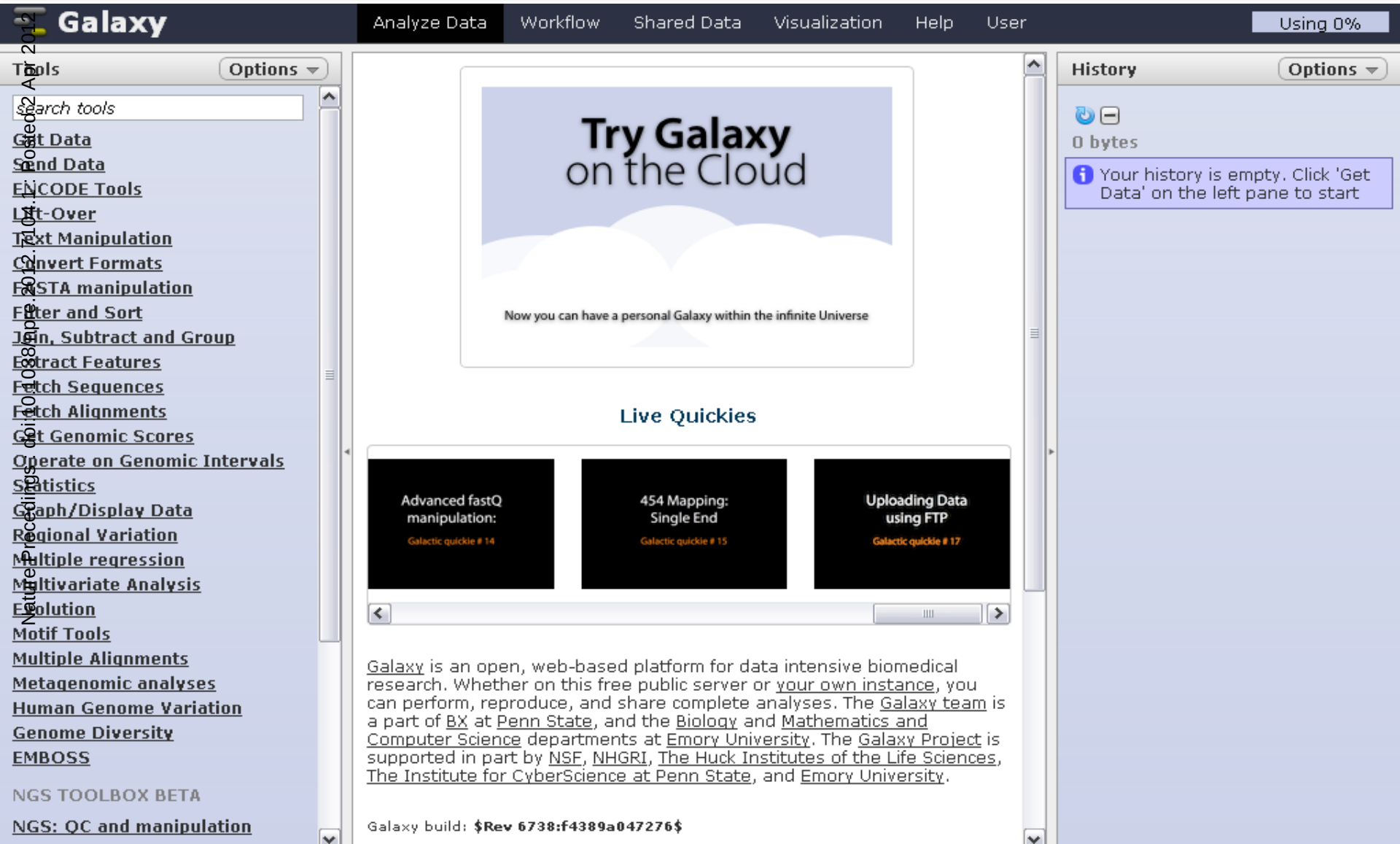
• Next generation sequencing on Amazon cloud • Taverna-Galaxy integration • CDK plugin for cheminformatics • Taverna 3 OSGi • SCUFL2 workflow bundle language • Taverna infrastructure VMs

averna is an open source and domain-independent Workflow anagement System – a suite of tools used to design and execute scientific workflows and aid *in silico* experimentation.

Taverna has been created by the myGrid team and funded through the OMII-UK. The project has guaranteed funding till 2014.

The Taverna suite is written in Java and includes the **Taverna Engine** (used for enacting workflows) that powers both the **Taverna Workbench** (the desktop client application) and the **Taverna Server** (which allows remote execution of workflows). Taverna is also available as a **Command**

Galaxy



Nature Precedings : doi:10.1038/npre.2012.7104.1 : Posted 2 Apr 2012

Emerging efforts for data management

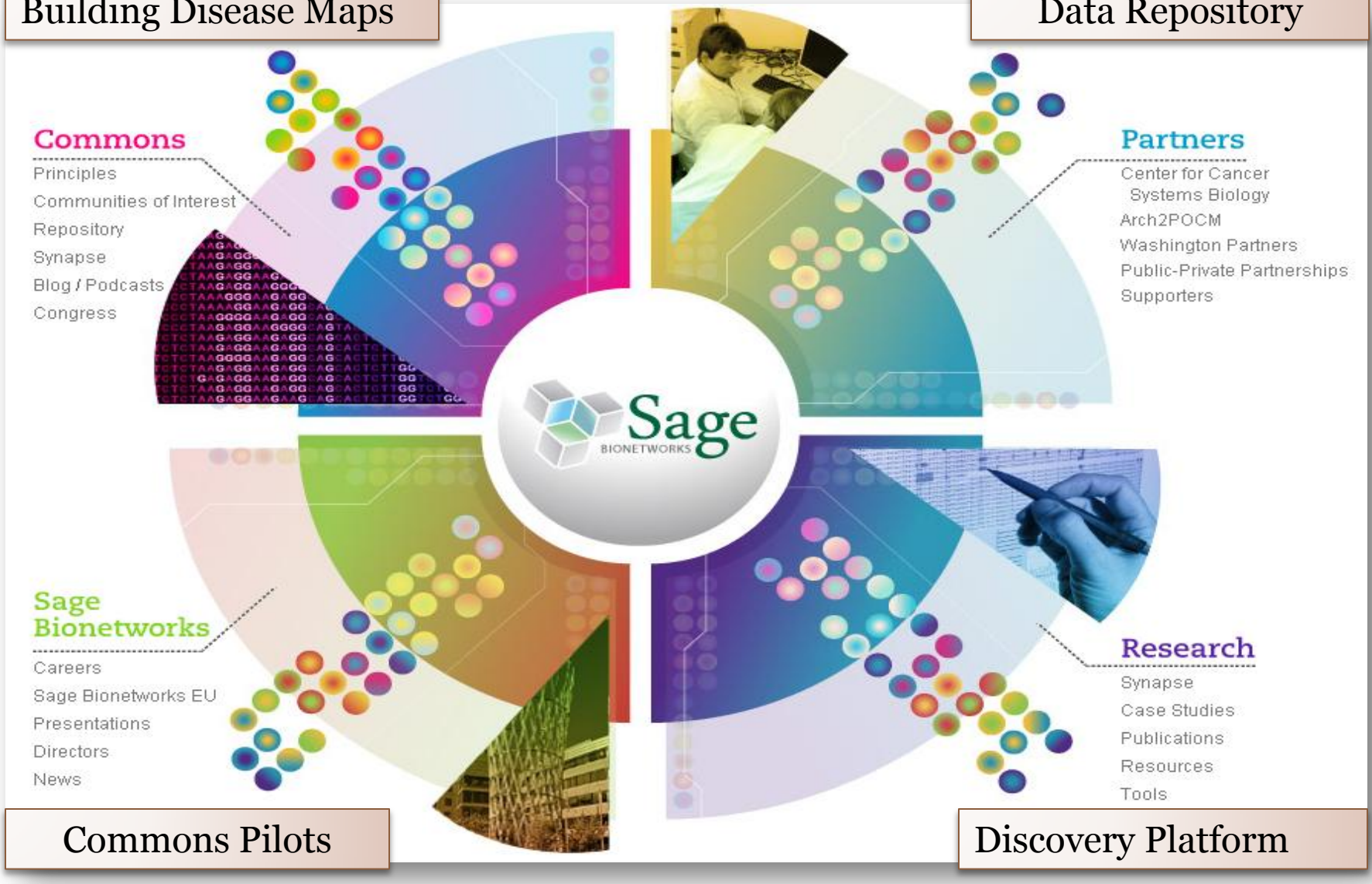
- **Sage Bionetworks**
  - Sage Bionetworks is currently focused on establishing a platform for data acquisition and curation.
  - The future aim of this platform is for modelling
  - It offers Open collaborative
  - Data for drug discovery

- **ELIXIR**
  - ELIXIR is a European effort that plans to build a biological data management infrastructure.

Saga bionetwork

Building Disease Maps

Data Repository



Commons Pilots

Discovery Platform

ELIXIR

about | for funders | for researchers | for industry | events | news & media | prep phase

Welcome to ELIXIR

"ELIXIR unites Europe's leading life science organisations in managing and safeguarding the massive amounts of data being generated every day by publicly funded research. It is a pan-European research infrastructure for biological information. ELIXIR will provide the facilities necessary for life science researchers - from bench biologists to cheminformaticians - to make the most of our rapidly growing store of information about living systems, which is the foundation on which our understanding of life is built."

Prof. Janet Thornton
ELIXIR Coordinator
Director of the EMBL-European Bioinformatics Institute

Sir Paul Nurse :

Latest news

BioMedBridges project began on 1 January 2012

BioMedBridges is an FP7-funded cluster project that will bring together the ESFRI Research Infrastructures in the Bio and Medical Sciences (BMS). BioMedBridges formally started on 1 January 2012 and over the duration of the four-year of the project will play a major role in providing

Contact us

ELIXIR Office
EMBL-European Bioinformatics Institute
Wellcome Trust Genome Campus
Hinxton, Cambridgeshire
CB10 1SD, United Kingdom

+44 (0)1223 494-444
+44 (0)1223 494-468
elixirpm@ebi.ac.uk

Software for systems biology

- Data handling
- **Network inference**
- Deep curation
- Dynamical simulation
- Model analysis

Data-driven network inference

- **Data**

  Multi-dimensional data require constructing probabilistic, causal gene networks

  - Genome scale DNA variation data
  - Gene expression data
  - Protein–protein interaction data
  - DNA–protein binding data
  - Complex binding data

Modelling

- A specific kind of modelling from large-scale data, known as **data-driven network-based modelling**

- The models known as inference networks, co-expression networks or association networks.

How data generate the model

High-throughput and time course experimental data

Use computational algorithms

Infer causal relationships among molecular entities (such as genes, transcription factors, proteins and metabolites)

Approaches to network inference models

- Based **on Bayesian inference techniques**

  Computing the probability of a hypothesis (i.e. relationship between two molecular entities) based on some kind of evidence or observations

Alternative techniques

- Regression methods
- Correlation methods
- Mutual information approaches

Mutual information approaches

- It is a dimensionless quantity that measures the extent to which one random variable is informative about another variable
- Zero mutual information → Independent
- Software tools -
  - R
  - MATLAB
  - BANJO

Standards in data-driven inference

- No true benchmarking standards
- Currently have their accuracy evaluated using simulated data

Recent efforts towards community-driven standardization initiated

- Sage Bionetworks
- Dialogue for Reverse Engineering Assessments and Methods (DREAM) initiative

DREAM initiative

Dialogue for Reverse Engineering Assessments and Methods

DREAM6 (2011)

The DREAM6 Challenges have been posted at

http://the-dream-project.org

Important Dates
Updated 6/03/2011

- June 1st, 2011 DREAM6 Challenges posted
- August 22nd, 2011 Prediction Submission DEADLINE
- September 12th, 2011 Scores Released, Best Performers Notified
- October 14th-19th, 2011 RECOMB Systems Biology / Regulatory Genomics / DREAM6 Conference (The DREAM6 conference track is on October 14th.)

Overview

DREAM is a Dialogue for Reverse Engineering Assessments and Methods. The main objective is to catalyze the interaction between experiment and theory in the area of cellular network inference and quantitative model building in systems biology. The fundamental questions for DREAM are simple: How can we assess how well we are describing the networks of interacting molecules that underlie biological systems? and How can we know how well we are predicting the outcome of previously unseen experiments from our models? The answer to these questions are not so simple. Researchers have used a variety of algorithms to deduce the structure of biological networks and/or to predict the outcome of perturbations to their systems. They have also eveluated the success success of their methodologies using diverse non-standardized metrics. However what is still needed, and what DREAM aims to achieve, is a fair comparison of the strengths and weaknesses of the methods and a clear sense of the reliability of the models that researchers create.

Nature Precedings : doi:10.1038/npre.2012.7104.1 : Posted 2 Apr 2012

Software for systems biology

- Data handling
- Network inference
- **Deep curation**
- Dynamical simulation
- Model analysis

Deep curation

- An alternative to data-driven network inference
- The deep curation approach creates a detailed molecular interaction map by information from publications, databases and high-throughput data

Comparison

Data-driven network inference	Deep curation
• Hypotheses about interactions are generated **automatically**	• deep curation approach constructs the model **manually** or **semi-manually**
• **Own hypotheses** can **not added or** not easy	• It easier for researchers to add their **own hypotheses** into it.
• **Not** give **rationale** to support the hypotheses	• Give **rationale** to support the hypotheses
• They do **not** provide **mechanistic details**	• They provide **mechanistic details**

It would be ideal to combine deep curation and data-driven approaches.

Resources

- Pathway databases provide information that can be used to create an initial draft of the pathway model.
  - Kyoto Encyclopedia of Genes and Genomes (KEGG)
  - Reactome
  - Panther pathway database
  - Pathway Commons
  - BioCyc

Kyoto Encyclopedia of Genes and Genomes (KEGG)

KEGG Home
Release notes
Current statistics
Plea from KEGG

KEGG Database
KEGG overview
Searching KEGG
KEGG mapping
Color codes

KEGG Objects
Pathway maps
Brite hierarchies

KEGG Software
KegTools
KEGG API
KGML

KEGG FTP
Subscription

GenomeNet

DBGET/LinkDB

Feedback

Kanehisa Labs

KEGG: Kyoto Encyclopedia of Genes and Genomes

Announcement
The KEGG website at www.kegg.jp has become the primary site of the KEGG database developed by Kanehisa Laboratories (see the article in the NAR 2012 Database Issue). The GenomeNet website at www.genome.jp operated by Kyoto University Bioinformatics Center will continue to mirror the KEGG database and provide additional KEGG-based analysis services (see Release notes).

Main entry point to the KEGG web service

KEGG2	KEGG Table of Contents	Update notes

Data-oriented entry points

KEGG PATHWAY	KEGG pathway maps [Pathway list]
KEGG BRITE	BRITE functional hierarchies [Brite list]
KEGG MODULE	KEGG modules [Module list]
KEGG DISEASE	Human diseases [Cancer | Infectious disease]
KEGG DRUG	Drugs [ATC drug classification]
KEGG ORTHOLOGY	Ortholog groups [KO system]
KEGG GENOME	Genomes [KEGG organisms]
KEGG GENES	Genes and proteins Release history
KEGG LIGAND	Chemical information [Compound classification]

Entry point for wider society

KEGG MEDICUS	Health-related information resource

Organism-specific entry points

KEGG Organisms — Enter org code(s) [ ] Go hsa hsa eco

Analysis tools

KEGG Mapper	KEGG PATHWAY/BRITE/MODULE mapping tools
KEGG Atlas	Navigation tool to explore KEGG global maps
KAAS	KEGG automatic annotation server
BLAST/FASTA	Sequence similarity search
SIMCOMP	Chemical structure similarity search
PathPred	Biodegradation/biosynthesis pathway prediction

REACTOME

Nature Precedings : doi:10.1038/npre.2012.7104.1 : Posted 2 Apr 2012

Search examples...

Browse Pathways

Map IDs to Pathways

Compare Species

Analyse Expression Data

If you would prefer to use our old website, click here.

Download

The following links allow you to download Reactome data in various formats:

- BioPax
- SBML
- Textbook
- Other formats

Try this

Have you got a set of genes or proteins, where you would like to understand the biological context

About Reactome

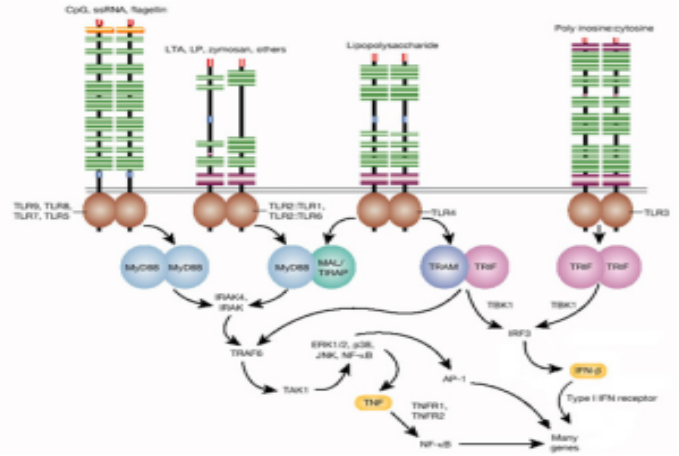
REACTOME is an open-source, open access, manually curated and peer-reviewed pathway database. Pathway annotations are authored by expert biologists, in collaboration with Reactome editorial staff and cross-referenced to many bioinformatics databases. These include NCBI Entrez Gene, Ensembl and UniProt databases, the UCSC and HapMap Genome Browsers, the KEGG Compound and ChEBI small molecule databases, PubMed, and Gene Ontology. ... [more]

Reactome Milestone

Reactome has achieved its milestone of curating reactions and pathways involving at least 5000 distinct human proteins... [more]

Tutorial

Pathway of the Month: Toll Receptor Cascades



Panther pathway database

PANTHER Classification System

LOGIN REGISTER CONTACT US HELP

Home | Browse | Genes and orthologs | Trees and HMMs | Pathways | Ontologies | Tools | Workspace

Community Pathway Curation | Browse Pathways | Search Pathways | Pathway Resources |

Search

Pathways

Go

Quick links

- Whole genome function views
- Gene expression tools
- cSNP tools
- Upload multiple gene IDs
- Community Curation
- My Workspace
- HMM scoring
- Downloads
- Genome statistics
- Site map

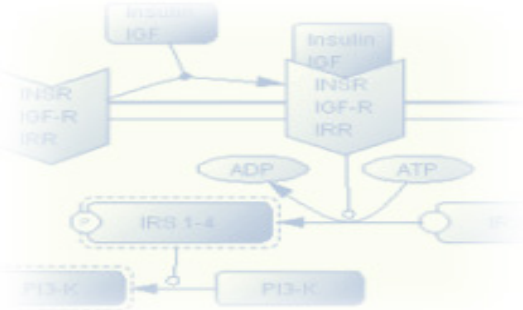
Newsletter subscription

Enter your Email:

Subscribe

PATHWAYS

PANTHER Pathway consists of over 165, primarily signaling, pathways, each with subfamilies and protein sequences mapped to individual pathway components. A component is usually a single protein in a given organism, but multiple proteins can sometimes play the same role. Pathways are drawn using CellDesigner software, capturing molecular level events in both signaling and metabolic pathways, and can be exported in SBML format. The SBGN view of the diagram can also be exported. Pathway diagrams are interactive and include tools for visualizing gene expression data in the context of the diagrams.
Pathway release calendar



NEW! Community Pathway Curation
Step-by-step instructions for the creation of structured pathways that can be published in the PANTHER Pathway database.

Browse Pathways
Browse a complete list of over 165 PANTHER pathways, pathway components and view diagrams.

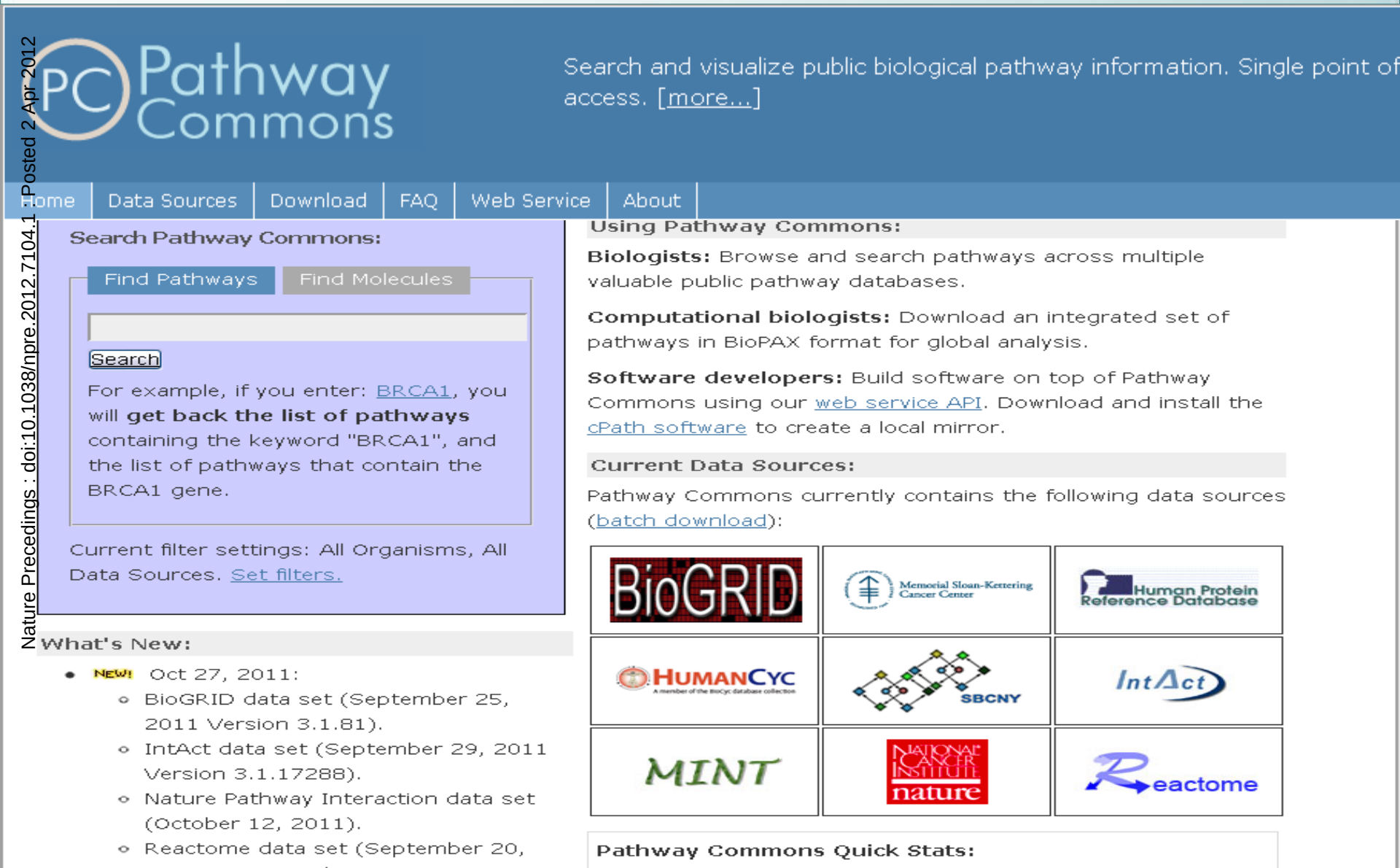
Search Pathways
Search pathways and pathway components by keywords such as molecular function, biological process, pathway relationships and interactions, or PANTHER families.

Pathway Resources
See a list of useful resources including links to other pathway datasets and tools for editing and viewing biological models.

List of Pathways
See a complete list of curated pathways.

Download pathway data

Pathway Commons



BioCyc

BioCyc
Database Collection

News

BioCyc version 16.0 contains 1763 genomes. Read more.

Information

Introduction to BioCyc
Guide to BioCyc
Update History
Webinars
1763 Databases
Guided Tour
Pathway Tools Software
Publications
Linking to BioCyc
External Links
Contact Us

Services

Subscribe to BioCyc
Metabolic Posters
Genome Posters
Software/Database Downloads
Registry
Web Services

ABOUT BIOCYC

BioCyc is a collection of 1763 Pathway/Genome Databases. Each database in the BioCyc collection describes the genome and metabolic pathways of a single organism.

New to BioCyc? Typical usage:

- Select a database (genome) to search by clicking "change organism database" at top right
- Enter a gene name or pathway name in the box at top right and click Quick Search

Windows users: We strongly suggest you use Firefox instead of Internet Explorer to interact with this web site [more].
To learn more about BioCyc, read the Introduction to BioCyc or watch our instructional videos.

BIOCYC TOOLS

The BioCyc Web site contains many tools for navigating and analyzing these databases, and for analyzing omics data, including the following.

- Genome browser
- Display of individual metabolic pathways, and of full metabolic maps
- Visual analysis of user-supplied omics datasets by painting onto metabolic map, regulatory map, and genome map
- Comparative analysis tools

The downloadable version of BioCyc that includes the Pathway Tools software provides more speed and power than the BioCyc Web site [more]. Multiple database configurations are available for installation with the software including multiple *E. coli* and *Shigella* genomes, multiple *Bacillus* genomes, multiple *Mycobacterium* genomes, and multiple mammalian genomes.

BIOCYC PATHWAY/GENOME DATABASES

The BioCyc databases are divided into three tiers, based on their quality.

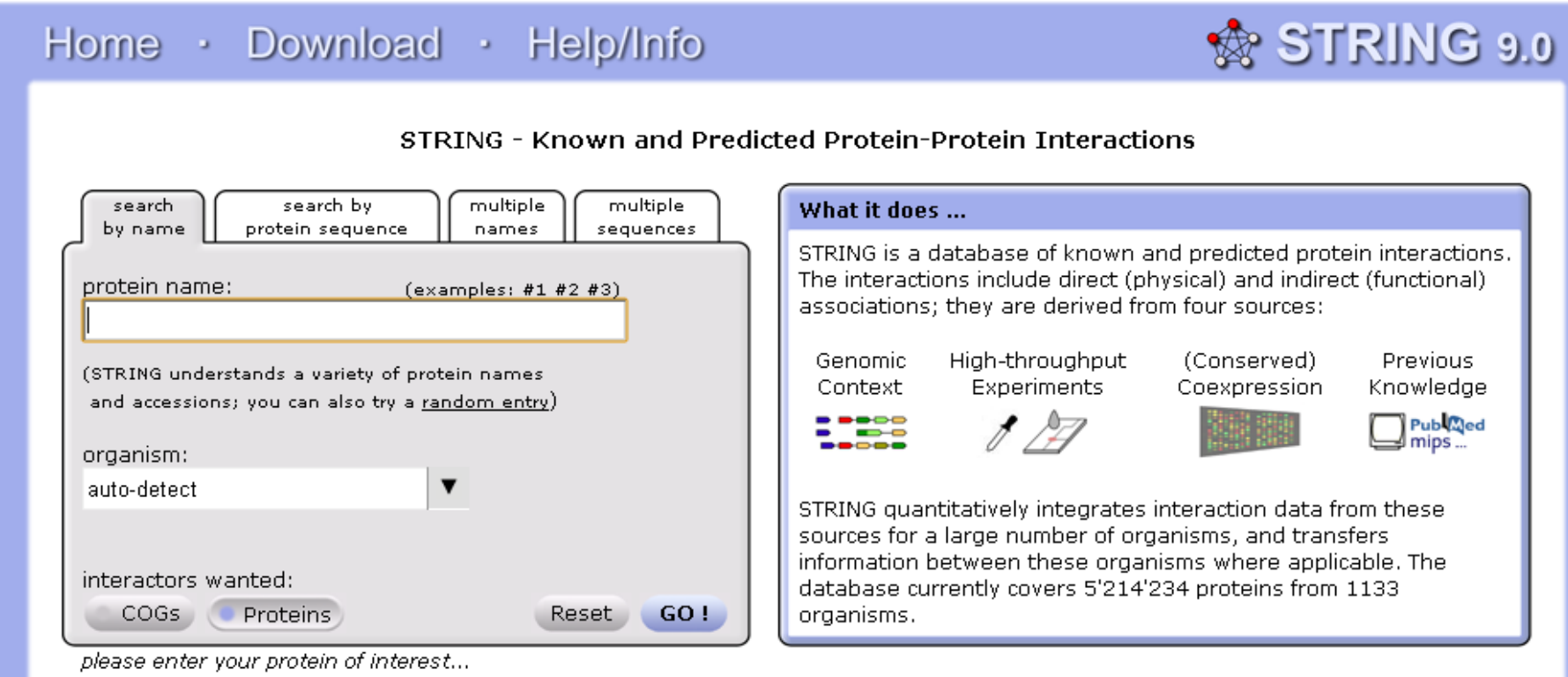
Tier 1 databases have received person-decades of literature-based curation, and are the most accurate. Tier 2 and Tier 3 databases contain computationally predicted metabolic pathways, predictions as to which genes code for missing enzymes in metabolic pathways,

Meta-databases

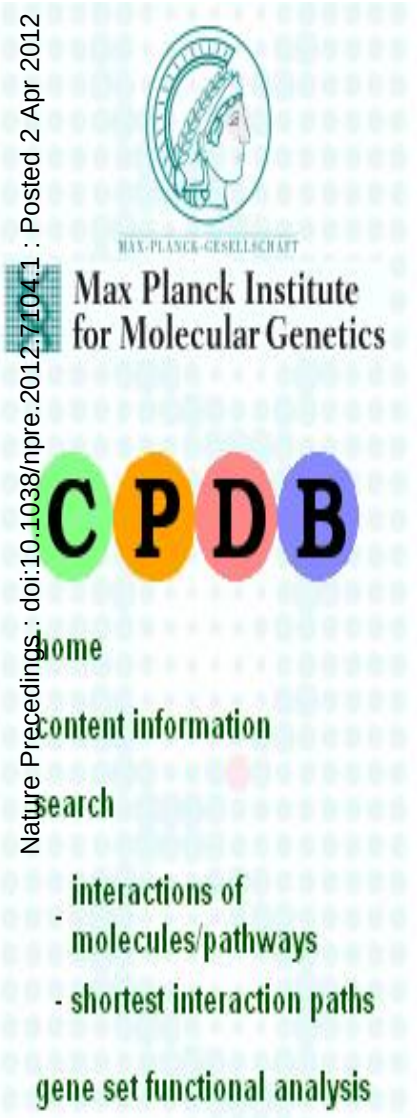
- Search Tool for the Retrieval of Interacting Genes/Proteins (STRING)
- ConsensusPathDB (CPDB)
  Max-Planck-Institute for Molecular Genetics

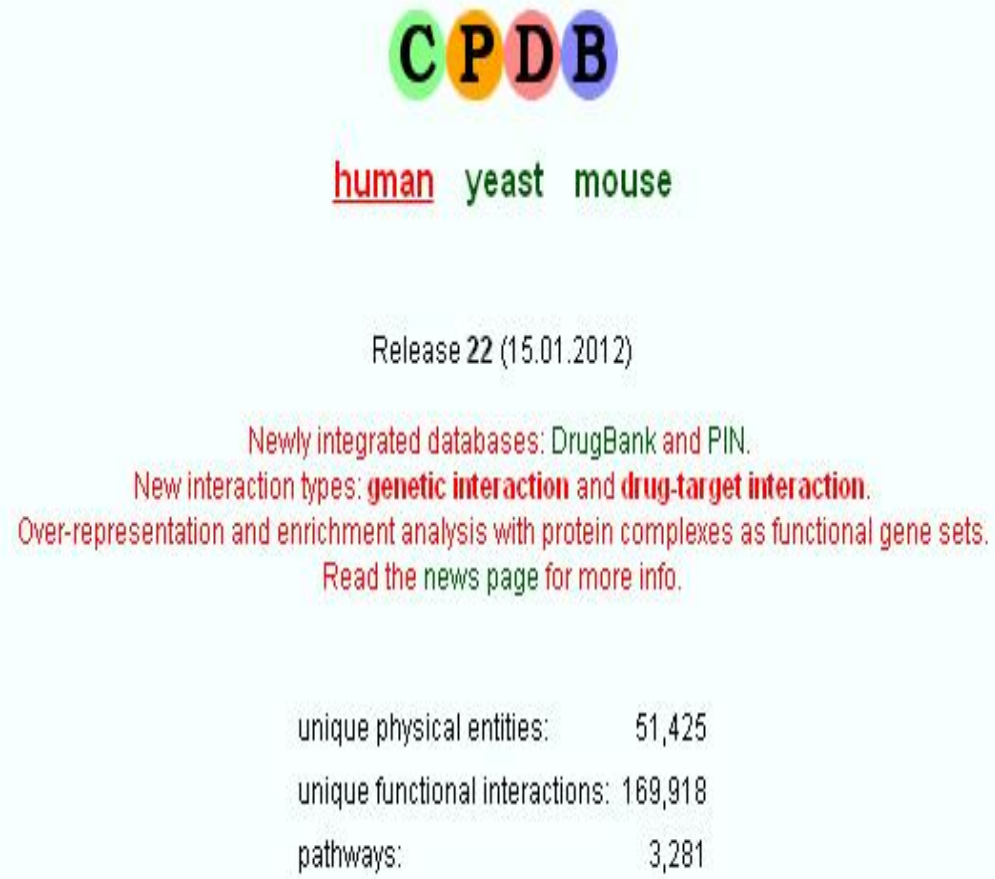
Search Tool for the Retrieval of Interacting Genes/Proteins (STRING)

Nature Precedings : doi:10.1038/npre.2012.7104.1 : Posted 2 Apr 2012



ConsensusPathDB (CPDB)





CPDB ConsensusPathDB

- Reactome
- Kegg
- Humancyc
- Pid
- Biocarta
- Netpath
- Inoh
- Ehmn
- Phosphositeplus
- Innatedb
- Intact-ss
- Intact-ls
- Dip
- Mint
- Hprd
- Corum
- Biogrid
- Mips-mppi
- Bind
- Matrixdb
- Pin
- Spike
- Pig
- Phosphopoint
- Pdzbase
- Pharmgkb
- Smpdb
- Signalink
- Drugbank
- Wikipathways

Machine-readable and model- representation standards

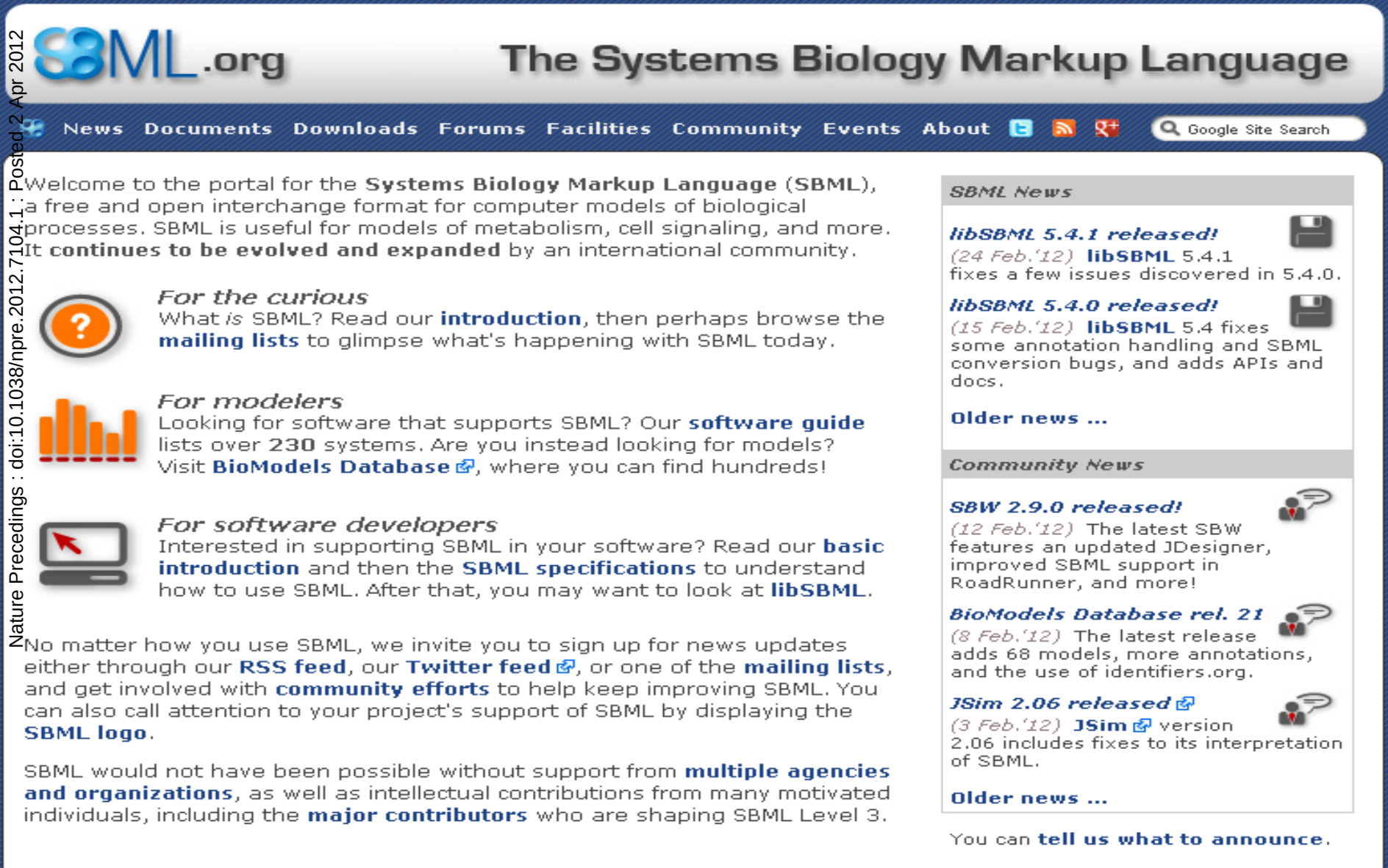
- Systems Biology Markup Language (**SBML**)
- Biological Pathways exchange (**BioPAX**)

Both were designed to represent biomolecular networks

Systems Biology Graphical Notation (**SBGN**)

Designed to standardize a human-readable pathway notation

SBML



BioPAX

BioPAX.org

home | specification | documentation | wiki | validator

BioPAX - Biological Pathway Exchange

BioPAX is a standard language that aims to enable integration, exchange, visualization and analysis of biological pathway data. Specifically, BioPAX supports data exchange between pathway data groups and thus reduces the complexity of interchange between data formats by providing an accepted standard format for pathway data. By offering a standard, with well-defined semantics for pathway representation, BioPAX allows pathway databases and software to interact more efficiently. In addition, BioPAX enables the development of pathway visualization from databases and facilitates analysis of experimentally generated data through combination with prior knowledge. The BioPAX effort is coordinated closely with that of other pathway related standards initiatives namely; PSI-MI, SBML, CellML, and SBGN in order to deliver a compatible standard in the areas where they overlap.

Status and Development

BioPAX Level 3 was finalized by the community in November 2009 and covers metabolic pathways, molecular interactions, signaling pathways (including molecular states and generics), gene regulation and genetic interactions. The current release of the BioPAX Level 3 can be found in the links below and the most up to date version is available to download from sourceforge.

- http://www.biopax.org/release/biopax-level3.owl
- http://www.biopax.org/release/biopax-level3-documentation.pdf

- Home
- Specifications
- Documentation
- Proposals and Workgroups
- Contributions, Editors and SAB
- Validation
- PaxTools
- Archive

BioPAX paper in Nature Biotechnology

- **Special Section**

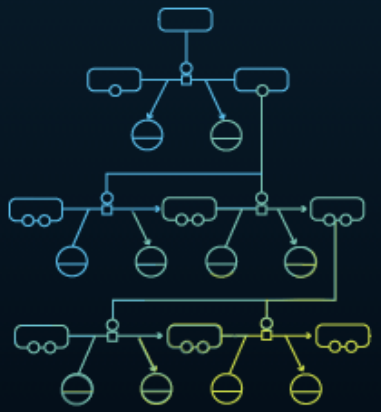
Perspective by Demir *et al.* **FREE**

Incompatible data storage formats have hindered the sharing and analyses of digital representations of biological pathways. BioPAX is a standardized

SBGN

SBGN Systems Biology Graphical Notation

Home | News | Documents | Lists | Community | Events | About | Google Site Search...



A Visual Notation for Network Diagrams in Biology

Welcome to the global portal for documentation, news, and other information about the Systems Biology Graphical Notation (SBGN) project, an effort to standardize the graphical notation used in maps of biological processes.

The following links will bring you to the more popular pages of the website. Much more is available though. Explore using the top menu.

Specifications	Software	Examples
Next meeting	Discussion	Cite
SBGN Competition 2011		

SBGN is the work of many people. It would not have been possible without the generous support of multiple organizations over the years, for which we are very thankful.

SBGN News

(10 Dec. '11) The Milestone 2 of libSBGN has been released.

Rules for model annotation - MIRIAM

BioModels.Net

Home | Database | MIRIAM | SBO | MIASE | KiSAO | TEDDY | Qualifiers | Events | Contact

MIRIAM

Table of content

1. MIRIAM Standard
2. MIRIAM URIs
3. MIRIAM Registry
4. Publications

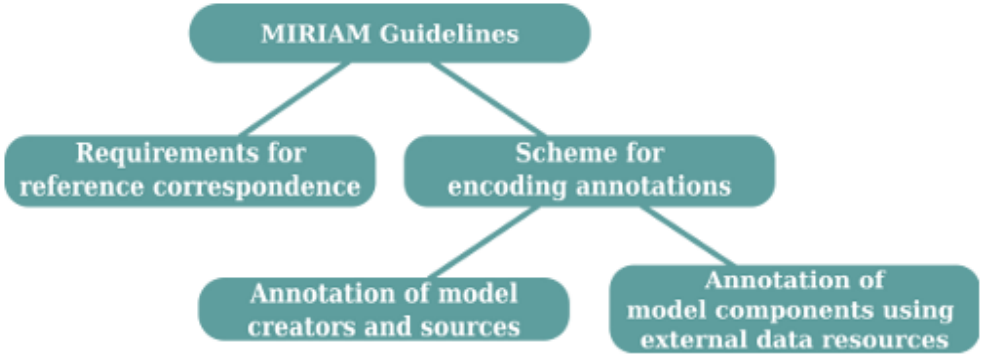
MIRIAM is an effort to standardise the ***Minimal Information Required In the Annotation of Models***, so that different groups can collaborate on annotating and curating computational models in biology. The goal of the project, initiated by the **BioModels.net** effort is to produce a set of guidelines suitable for use with any structured format for computational models.

MIRIAM is a registered project of the **MIBBI** (Minimum Information for Biological and Biomedical Investigations).

If you are looking for the online resources providing support for the MIRIAM Standard, please go to: **MIRIAM Registry**.

MIRIAM Standard

The *MIRIAM Standard* is composed of three parts: **reference correspondence**, **attribution annotation**, and **external resource annotation**. These are described below:



Tools and model databases for support deep curation

- CellDesigner

Plug-in application programming interface (API) for CellDesigner

- **Edinburgh Pathway Editor** (EPE)
- **Jdesigner**

  Provide graphical editing and visualization capabilities

- **PathVISIO** (for pathway curation)
- **Cytoscape** (which is a widely used tool for the visualization of molecular networks)

CellDesigner

CellDesigner.org

home | features | downloads | documents | models | forum | news | links | contacts

home

Menu

- Features
- Download
- Quick Tutorial
- Documents
- Help
- Simulation
- Model Repository
- Plugins
- News
- Links

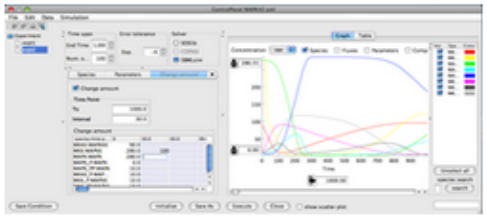
Related Link

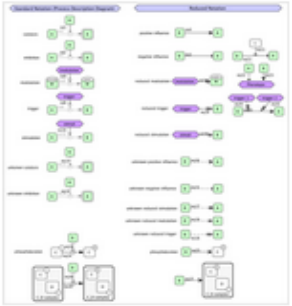
- systems-biology.org
- The Systems Biology Institute
- Kitano Symbiotic Systems Project
- PANTHER database
- BioModels.Net

CellDesigner™: A modeling tool of biochemical networks

Download CellDesigner

Current Release Version: CellDesigner 4.2
supports **Reduced Notation, Simulation Parameter Polymorphism, SBMLsim solver**... and more. find out more...





For 64bit Windows Users (2010/03/04)
CellDesigner support 32 bit only. Please install CellDesigner onto 32 bit mode.

For 64bit Linux Users (2011/01/26)
CellDesigner support 32 bit only. Please install CellDesigner onto 32 bit mode.
$ sudo aptitude install ia32-libs

*** How to install CellDesigner4.2 on 64bit Linux (2011/10/15)**

1. Install "LSB (Linux Standard Bas)" if not installed on your system
2. Install "ia32-libs" if not installed on your system
3. Launch CellDesigner4.2-Installer with following command:
   $ ./CellDesigner-4.2-linux-installer.bin --mode xwindow

Check also:

- **Plugins** / **Utilities**
- BioModels.net models simulation results with CellDesigner 4.0

Headlines

Plugin: BioPAXLevel 3 Export plugin is now available (2011/11/8)

***CellDesigner 4.2** is now available* (2011/10/05)
Ver.4.2 supports **Reduced Notation, Simulation Parameter Polymorphism, SBMLsim solver**... and more versionup info...

Tool: iPathways 1.2 (IAppli) released (2011/08/30)

Plugins: SBSI CellDesigner plugin. Check Plugins page...(2011/1/19)

Plugins: .Merge Models & Payao Uploader. Check Plugins page...(2010/10/10)

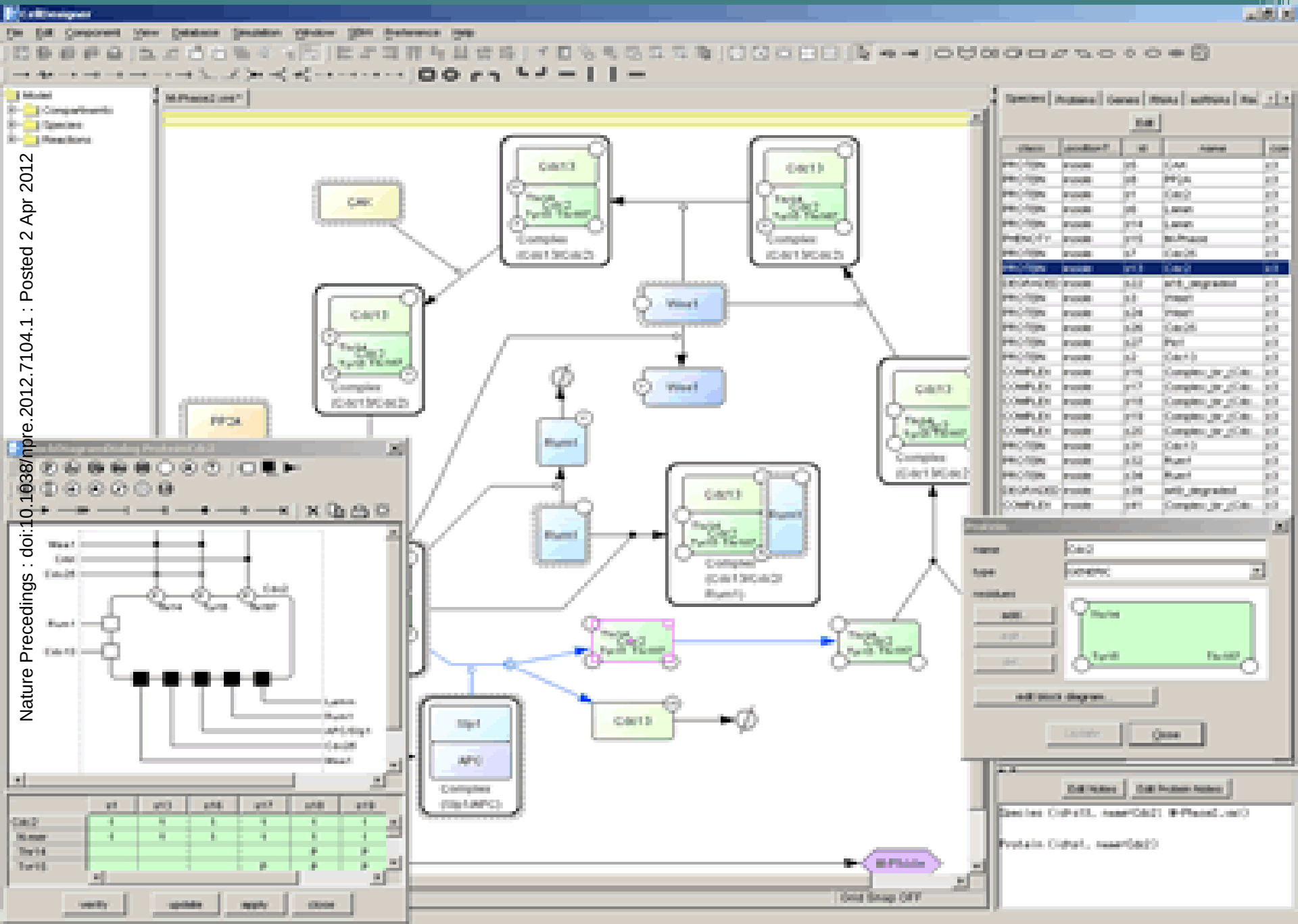
CellDesigner4.1 Tutorial at **ICSB 2010,** Oct 10th, 2010.

ICSB 2010 Tutorial "Community Pathway Curation: Tools and Technologies" also features CellDesigner for curation.

Tutorial for Simulation (by Dr. Richard Adams) materials available **here** (2010/08/12)

Plugin: PathwayAccess: CellDesigner Plugins for Pathway Databases. Check

Nature Precedings : doi:10.1038/npre.2012.7104.1 : Posted 2 Apr 2012



Plug-in application for CellDesigner

Plugins	
BioPax Export	Export Model data to BioPAX Level3 format
Import Notes	Import Notes data from CSV file
Mapping Array Mass	Mapping experimental array data to the model
Merge Models	Merge CellDesigner's models
PathwayAccess	a suite of CellDesigner plugins directly interact with pathway datasources to download and integrate one or more pathways to a CellDesigner model, and upload (or commit) a CellDesigner model to a datasource.
Payao Uploader	Upload CellDesigner model to Payao system
SBML2SMW	A CellDesigner plugin and a semantic translation server to bidirectional read/write from/in Semantic Media Wiki.
SBMLsqueezer	A CellDesigner plugin to generate kinetic rate equations for biochemical networks
SBSI CellDesigner plugin	A CellDesigner plugin to bridge the functionalities of CellDesigner and SBSI
CellDesginer related utilities:	
BiNoM (Cytoscape Plugin)	Export CellDesigner maps to other formats, with partial support of CellDesigner semantics. Our Cytoscape plugin BiNoM can convert CellDesigner files to Cytoscape, and to BioPAX v.2.0 formats.
BioPP	BioPP suite is a web-based application that allows pathway knowledgebases stored in CellDesigner-SBML to be web published with an easily navigated user interface.
Payao	Payao is a community-based, collaborative web service platform for gene-regulatory and biochemical pathway model curation. Payao reads the models in SBML format, displays them with CellDesigner, and provides an interface for model enrichment (adding tags and comments to the models) for the access-controlled community members.
CellPublisher	CellPublisher is a free and open web server to share CellDesigner diagrams with a wide audience.

Challenges of deep curation

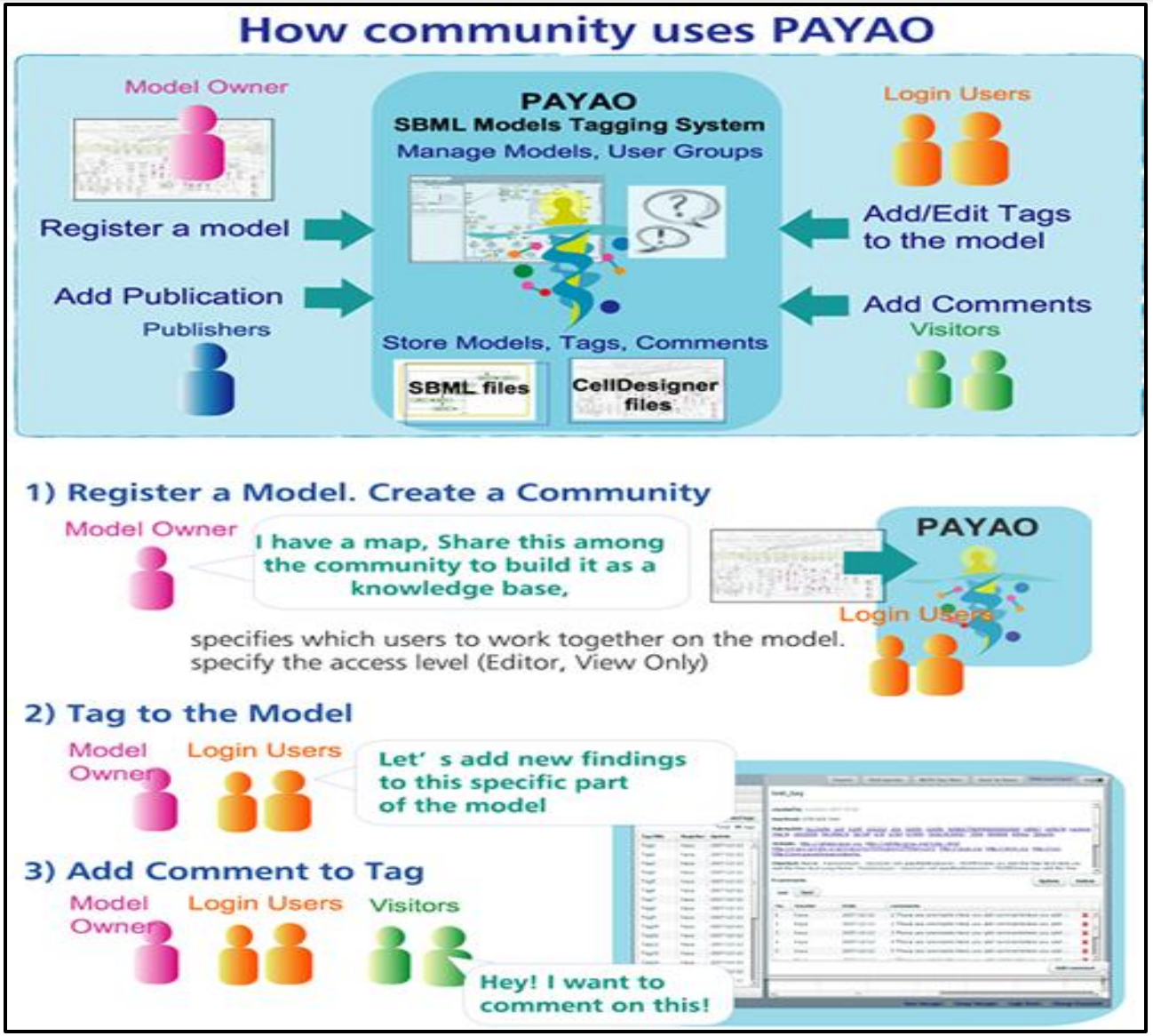
- The quality of pathways is often compromised by **fragmentation** and **inaccuracy**
- **Gold standard** -use curated maps that have been carefully built by a small group of people who spend months studying a pathway
  - Epidermal growth factor receptor (EGFR) pathway
  - Toll-like receptor pathway
  - Mammalian target of rapamycin (mTOR) pathway
  - Yeast cell cycle
  - E2F pathway

- Pathways updation and validation
- Manually creating large-scale network maps from the literature is extremely labour-intensive
  - Automate knowledge discovery
  - Automated literature mining
  - collaborative curation
    - **Payao system**
    - **WikiPathways**

BUT……

Insufficient participation from active users

Payao system



Nature Precedings : doi:10.1038/npre.2012.7104.1 : Posted 2 Apr 2012

WikiPathways

page | discussion | view source | history

Log in / create account

search

navigation
- Home
- Help

pathway
- Create
- Browse
- Mish List
- Download
- Web service API

overview
- Recent Changes
- Most Viewed
- Most Edited
- New Pathways
- Statistics

community
- About us
- Contact us
- How to cite
- Curation events

Welcome to WikiPathways BETA

WikiPathways is an open, public platform dedicated to the curation of biological pathways by and for the scientific community. More about WikiPathways...

Finding Pathways

Search

Search

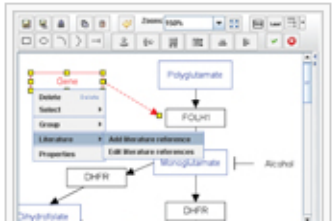
You can search by:
- Pathway name *(Apoptosis)*
- Gene or protein name *(p53)*
- Any page content *(cancer)*

Browse

Browse by species and category

Contributing New Pathways

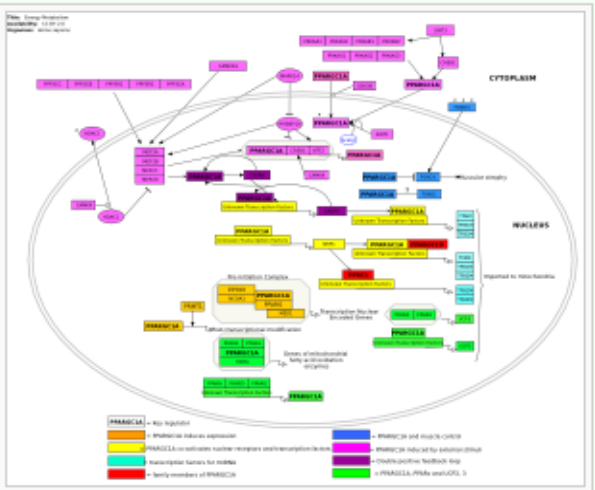
Create



Suggest

Today's Featured Pathway

Energy Metabolism (Homo sapiens)



Energy Metabolism

Latest discussions

27 February 2012

Suggested improvements to pathway (5) by Kristina Hanspers

18 February 2012

Welcome, new editor (1) by Alexander Pico

17 February 2012

Software for systems biology

- Data handling
- Network inference
- Deep curation
- **Dynamical simulation**
- Model analysis

In silico simulation models

- ***Building a model is like eating an elephant: it's hard to know where to begin.***

--J. W. Haefner

- Molecular interaction maps provide a static picture
- Dynamical simulations are mostly based on models created by the deep curation approach
- This is because deep curation captures
  - **Causality**
  - **Stoichiometry**
  - **Mechanisms of interactions**

Methods used in computational systems biology

- Systems of ordinary differential equations (ODEs)
- Petri nets
- Pi calculus
- Partial differential equations (PDEs)
- Cellular automata (CA) methods
- Agent-based systems
- Hybrid approaches

Standards for simulations

- Several standardization efforts empower the modelling community –
  - SBML
  - SBGN
  - MIRIAM

Language of Simulation

- ✓ **Simulation Experiment Description Markup Language (SED-ML)**
  - XML-based
- ✓ **Systems Biology Results Markup Language (SBRML)**
  - Complementary language to SBML

Tools for simulation

- **MATLAB**
- **Complex Pathway Simulator (COPASI)**
- **Systems Biology Workbench (SBW)**

  It is a software platform that allows multiple applications -

  - Modelling
  - Analysis
  - Visualization

Complex Pathway Simulator (COPASI)

COPASI

Download | Documentation | Forum | Screenshots | Meetings | Team | Funding

Search

in:

Entire Site | Go

Latest Versions

Stable:
COPASI 4.8 (Build 35)

Development:
COPASI 4.6.33 (Development)

Login

Login as...

User:

Password:

Remember me

Login

[Register | I forgot my pass]

COPASI: biochemical network simulator

COPASI is a software application for simulation and analysis of biochemical networks and their dynamics. COPASI is a stand-alone program that supports models in the SBML standard and can simulate their behavior using ODEs or Gillespie's stochastic simulation algorithm; arbitrary discrete events can be included in such simulations.

COPASI carries out several analyses of the network and its dynamics and has extensive support for parameter estimation and optimization. COPASI provides means to visualize data in customizable plots, histograms and animations of network diagrams. (list of features).

Language bindings for COPASI 4.8 (Build 35)

By: gauges on: Sat 24 of Dec., 2011 07:06 EST (451 Reads)

Language bindings for COPASI 4.8 (Build 35)

The download page for the language bindings should work now. Happy downloading and sorry for the delay.

Read More (1176 bytes)

COPASI 4.8 (Build 35) Released

By: Stefan Hoops on: Tue 20 of Dec., 2011 12:21 EST (903 Reads)

The COPASI team announces the immediate availability of the stable release COPASI 4.8 (Build 35).

Read More (1251 bytes)

Systems Biology Workbench (SBW)

Systems Biology Workbench (SBW)

SF project page | SBW Home | SBW Module Project | Search Site

SBW - Core Development

Systems Biology Workbench

- About
- How to Install
- SBW Project Page
- SBW Modules
- SBW Home
- Documentation

What is SBW

Researchers in quantitative systems biology make use of a large number of different software packages for modeling, analysis, visualization, and general data manipulation. The Systems Biology Workbench (SBW), is a software framework that allows heterogeneous application components-written in diverse programming languages and running on different platforms-to communicate and use each others' capabilities via a fast binary encoded-message system. Our goal was to create a simple, high performance, open-source software infrastructure which is easy to implement and understand. SBW enables applications (potentially running on separate, distributed computers) to communicate via a simple network protocol. The interfaces to the system are encapsulated in client-side libraries that we provide for different programming languages.

At the last count, there were over 75 different packages for simulating cellular networks (see www.sbml.org). This proliferation of tools has resulted in a variety of capabilities and interfaces. Though welcome in many respects, this proliferation has resulted in two unwelcome side effects:

1. Each tool uses its own format, often undocumented, to store models. The result is that a model saved in one tool cannot be loaded into another. This obviously hinders the free exchange of models from one tool to another.

2. The second problem is that many of the tools duplicate each other's capabilities. Writing simulation tools takes time, and many of the projects are short-lived, which means that the authors are unable to develop the tools very far. As a result, many of the tools provide similar functionality. Unlike other software development communities, there is little tradition of code reuse in the system biology community. As a result, the community has seen much duplicated effort.

Model Interchange The first problem, that of model exchange, has been addressed by introducing a standard format for all tool writers to employ. This standard is called Systems Biology Markup Language (SBML) Along with CellML (www.cellml.org), the introduction of a standard format is beginning to make a significant impact on tools writers, and the majority of the most widely used tools now employ SBML as a means to exchange models.

Tools support Petri net modelling

- **ePNK**
- **Time Petri Net Analyser (TINA)**
  a toolbox for the editing and analysis of Petri nets
- **WoPeD**
  a tool for modelling, simulation and analyses

ePNK

Informatics and Mathematical Modelling

Technical University of Denmark

DTU

Teaching | Research | Industrial collaboration | About IMM | News

[THE EPNK HOME PAGE]

Overview

The ePNK is a platform for Petri net tools based on the PNML transfer format. Its main idea is to provide generic Petri net types, which can be easily plugged into it, and to provide a simple generic GMF editor, which can be used for graphically editing nets of any plugged in type. Additional functionality can be plugged in. The ePNK is implemented based on the eclipse platform, and runs on all hardware platforms supporting eclipse (right now it is deployed for Eclipse Galileo and Helios).

The current version (June 17, 2011) of the ePNK is 0.9.2, which fully supports PNML and all the Petri net types defined in the International Standard ISO/IEC 15909-2. Though the current version is quite stable and supports many new and interesting features (see release notes), we do not call this version 1.0. The reason is that the graphical editor of the ePNK is still not fully exporting all graphical features of PNML (an incentive for me to implement some of these features would be that someone tells me about a concrete project or application where this would be really needed).

If you have any comments, suggestions, error reports, please send an email to Ekkart Kindler (eki@imm.dtu.dk).

Current version

New: As of June 17, version 0.9.2 of the ePNK is released, which runs on eclipse Helios (a version for Galileo exists too, but is not yet deployed). You will find more information on how to install the current version of the ePNK on your platform at the Installation Details page – there, you will also find an overview of the different versions and the functionality that was added for the different versions.

If you prefer it the quick way, here is the ePNK update site for Eclipse Helios:
http://www2.imm.dtu.dk/~eki/projects/ePNK/helios/update/ or look at the release notes of version 0.9.2.

Time Petri Net Analyser (TINA)

TINA

TIme petri Net Analyzer

Home | News | Download | Papers | Contact | Friends

What is Tina ?

Tina is a toolbox for the edition and analysis of **Petri Nets** and **Time Petri Nets**, developed in the OLC group of LAAS/CNRS. General Petri nets information can be found on the Petri Nets World site.

The Tina toolbox includes the tools:

nd (NetDraw): An editor for graphically or textually described Petri nets, Time Petri net and Automata. Interfaced with analysis tools below and drawing facilities.

tina: Construction of reachability graphs. Inputs nets in textual or graphical format. Outputs graphs in human readable form or in various formats for available model checkers and equivalence checkers.
This tool is described in [4] and [9]. Depending on options retained, it builds:

- The coverability graph of a Petri net, by the Karp and Miller technique.
- The marking graph of a bounded Petri net, checking boundedness on the fly.
- Partial marking graphs of a Petri net, by the covering steps methods of [6][7], the method of persistent sets, or several combinations of them explained in [8].
- Various state space abstractions for Time Petri nets (state class graphs), following the techniques introduced in [1][2] and further refined in [3][5]. Depending on option selected, the construction perserves markings, states, LTL properties, or CTL* properties of

WoPeD

WoPeD
Workflow Petri Net Designer

search | contact | imprint

Download WoPeD

Home
Team
Screenshots
Modeling
Download
Publications & Docs
Links

News

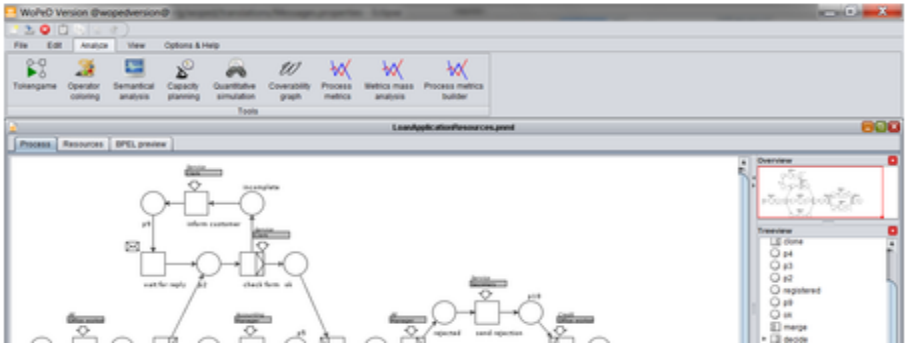
July 2011

WoPeD 2.6.0 has been released on July 5th, 2011!
The WoPeD → team is proud to present a new version of WoPeD which is ready to → download now!

→ See older News...

What is WoPeD?

WoPeD (**Wo**rkflow **Pe**tri Net **D**esigner) is an open-source software developed at the Cooperative State University Karlsruhe under the GNU Lesser General Public License (LGPL). The main goal of this tool is to provide an easy-to-use software tool for modelling, simulating and analyzing workflow process and resource descriptions using → workflow nets, an extended class of → Petri nets initially introduced by Wil van der Aalst (TU Eindhoven). WoPeD is mainly indended to be used by researchers, teaching staff or students dealing with the application of Petri nets to the area of workflow or business process management. WoPeD has already been successfully used as modelling tool in lectures and student assessment projects. WoPeD is maintained via Sourceforge, a web-based, open source development platform. The current development progress can be followed on the WoPeD project homepage at Sourceforge.



Software for systems biology

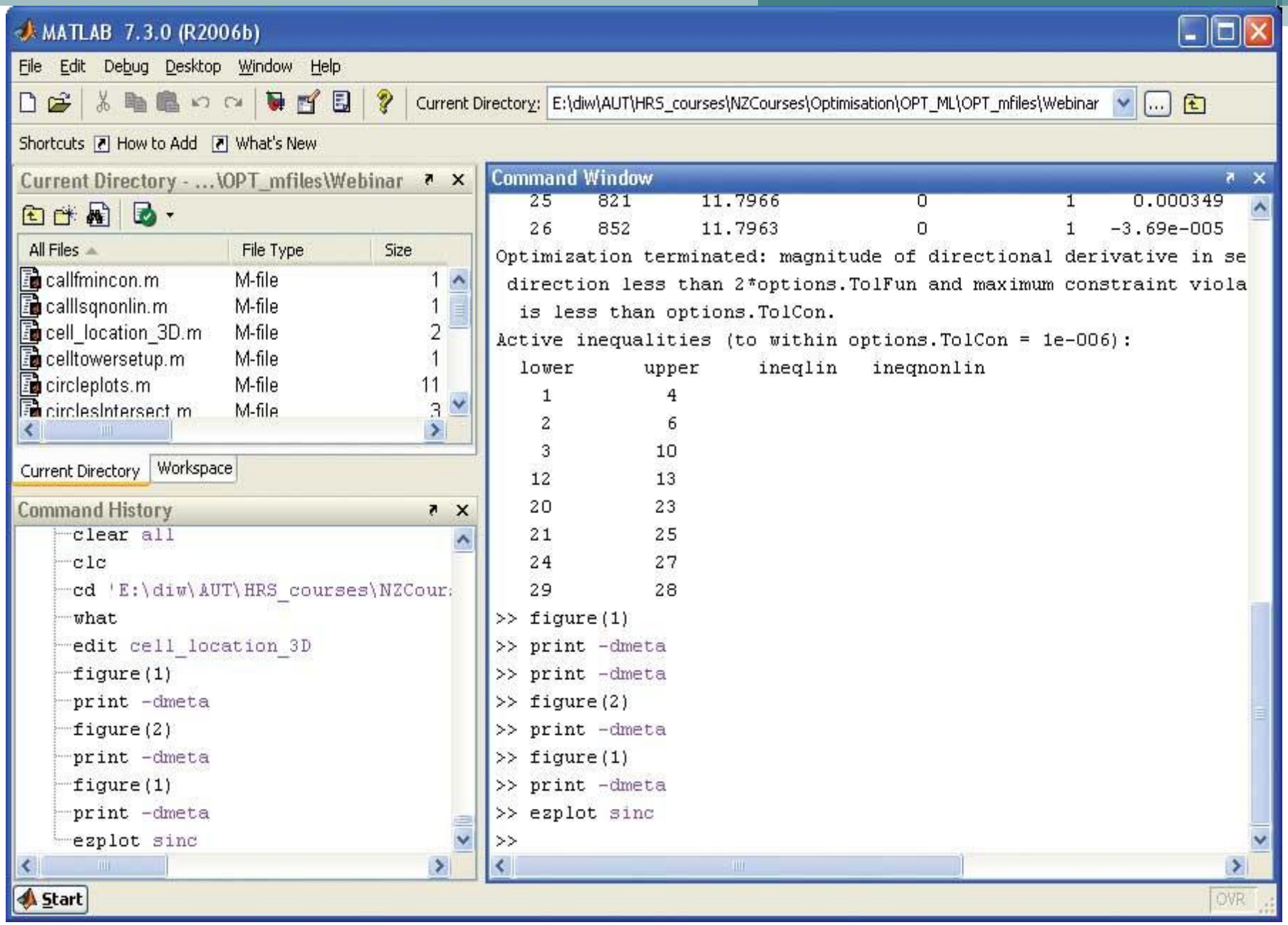
- Data handling
- Network inference
- Deep curation
- Dynamical simulation
- **Model analysis**

Model-analysis methods

- Sensitivity analysis
- Metabolic control analysis

Model-analysis tools

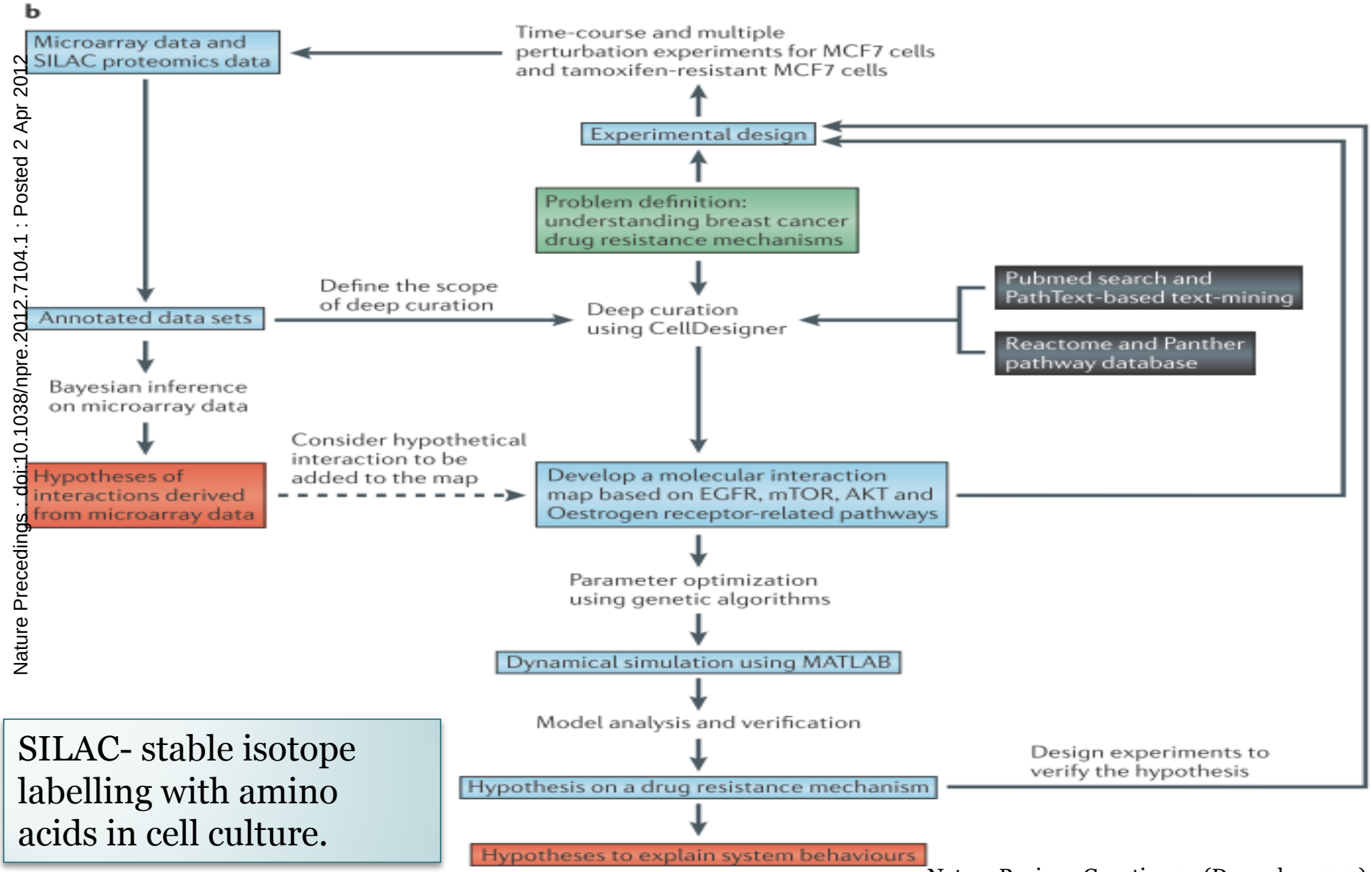
MATLAB



Sensitivity analysis

- SBML-SAT 
- MATLAB SimBiology MathWorks
- ByoDyn CBBL at GRIB-IMIM/UPF
- SensSB SensSB

Case study



SILAC- stable isotope labelling with amino acids in cell culture.

Nature Reviews Genetics, 12 (December 2011)

Conclusion

- There are lots of tools are available to Perform system biology but **integration of these tools** are necessary to increase the capability of scientist to explore more and more biological systems

- Require **innovative idea** to make this system biology more easy and comfortable for scientist

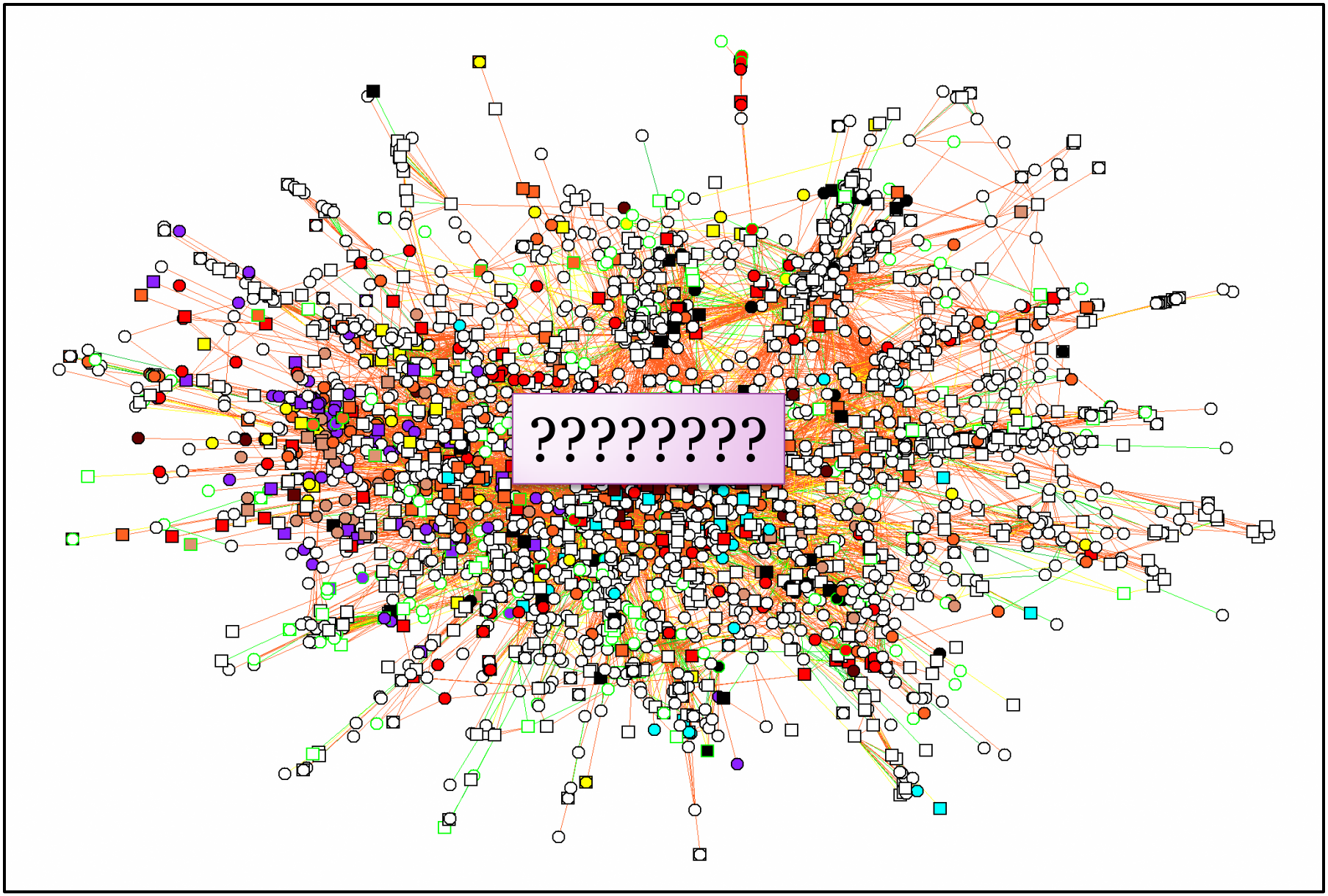
References

1. E.C. Butcher, E.L. Berg and E.J. Kunkel, Nature biotechnology, 22 (2004) 1253.

2.S. Ghosh, Y. Matsuoka, Y. Asai, K.Y. Hsin and H. Kitano, Nature Reviews Genetics, 12 (December 2011) 821.

3. Wu-tong WU (2006) Systems Biology in Drug Discovery. Pharmaceutical Biotechnology 11 (5):05

4. Cho CR, Labow M, Reinhardt M, van Oostrum J, Peitsch MC (2006) The application of systems biology to drug discovery. Current opinion in chemical biology 10 (4):294-302



Nature Precedings : doi:10.1038/npre.2012.7104.1 : Posted 2 Apr 2012